多模态古代汉语大语言模型 AI 九思 2.0 的设计与开发*

刘金柱[1] 王锦绣[1] 罗捷春[1] 李志芳[2,3] 袁方[1] 余静静[1] 龚丹[1] 谢雨霏[1] 罗婉滢[1] 郑苏楠[1] 陈旷心[1] 贺心雨[1] 张润哲[1] 夏婉婷[1] 谢佳延[1] 吕佳源[1] 吕萍[1] 余乐妍[1] 郑诗铭[1] 王金柳[1] 刘艺溶[4] 杨纯[5] 张曼丽[6] 吴翊嘉[1] 余锁湘[1] 汪靓[1] 刘根辉[1,7]

[1]（华中科技大学人文学院　武汉 430074）

[2]（武汉大学文学院 武汉 430062）

[3]（武汉大学古籍整理研究所　武汉 430062）

[4]（南京大学外国语学院 南京 210093）

[5]（成都市七中育才附属小学银杏校区 成都 610066）

[6]（安徽师范大学文学院 芜湖 241002）

[7]（华中科技大学铸牢中华民族共同体意识研究基地 武汉 430074）

摘要：[目的/意义] 随着生成式人工智能（AIGC）的快速发展，各类大模型由最初仅能处理单一文本模态的大语言模型，逐步升级为能够同时处理文本、图像、语音和视频等多模态数据的大语言模型。而国内面向古代汉语专业领域的大语言模型，仍主要聚焦于提升古汉语信息处理任务的性能，且以单一文本模态信息处理为主，在大语言模型的知识理解与问答交互能力，以及多模态信息处理方面，还有较大发展空间。基于此，华中科技大学全新推出了既掌握古汉语专业知识，又兼备古汉语应用能力，并支持多模态数据处理的古汉语多模态大语言模型——“AI 九思 2.0”，以为多模态古代汉语大语言模型的发展抛砖引玉。**[方法/过程]** 本文详细介绍了“AI 九思 2.0”的数据集构建、算力升级、模型训练、界面优化情况，并展示新版本“AI 九思”在古汉语语言知识和语言能力方面的表现。**[结果/结论]**全新升级的“AI 九思 2.0”在古代汉语文本理解及古代汉语知识问答领域展现出显著优势，且已经具备了一定的古文字（甲骨文、金文）图像理解能力，从而为推动古代汉语大语言模型的发展做出了应有的贡献。

关键词：AI 九思 2.0　古代汉语　大语言模型　多模态

分类号：G352

Abstract: Abstract: [Objective/Significance] With the rapid development of generative artificial intelligence (AIGC), various large models have gradually evolved from initially being capable of processing only single text modality in large language

*本研究受到国家社科基金重大项目“明代至民国汉语非韵书罕见同音类聚文献的音韵研究及数据库建设”（21&ZD297）、国家社科基金重大项目“草创时期甲骨文考释文献的整理与研究”（20&ZD307）、全国高等院校古籍整理研究工作委员会资助项目“孙奭《册府元龟》音义辑考”（批准编号：1835）、中央高校基本科研业务费“《册府元龟》语料库建设、整理与研究”（2020WKYXZX004）、中央高校基本科研业务费“《册府元龟》引书研究”（21WKFZZX016）、中央高校基本科研业务费专项资金项目“数字人文视域下《册府元龟》的深度整理与知识图谱构建研究”（YCJJ20242116）的资助。

作者简介：刘金柱，博士研究生；王锦绣，博士研究生；罗捷春，硕士研究生；李志芳，博士研究生；袁方，硕士研究生；余静静，硕士研究生；龚丹，硕士研究生；谢雨霏，硕士研究生；罗婉滢，硕士研究生；郑苏楠，硕士研究生；陈旷心，硕士研究生；贺心雨，硕士研究生；张润哲，硕士研究生；夏婉婷，硕士研究生；谢佳延，硕士研究生；吕佳源，硕士研究生；吕萍，硕士研究生；余乐妍，硕士研究生；郑诗铭，硕士研究生；王金柳，硕士研究生；刘艺溶，本科生；杨纯，硕士，小学教师；张曼丽，硕士研究生；吴翊嘉，本科生；余锁湘，本科生；汪靓，本科生；刘根辉，博士，教授，博士生导师，通信作者，E-mail：chnlab@mail.hust.edu.cn。

models to being able to handle multi-modal data such as text, images, voice, and video. However, domestic large language models targeting the ancient Chinese language field still mainly focus on improving the performance of ancient Chinese language processing tasks, and are mainly centered on single text modality information processing. There is still considerable room for development in terms of the knowledge understanding and question-answering interaction capabilities of large language models, as well as in multi-modal information processing. Based on this, Huazhong University of Science and Technology has newly launched the ancient Chinese multi-modal large language model “AI Jiusi 2.0”, which not only masters ancient Chinese professional knowledge but also possesses ancient Chinese application capabilities and supports multi-modal data processing, aiming to set an example for the development of multi-modal ancient Chinese large language models. **[Method/Process]** This paper details the dataset construction, computing power upgrade, model training, and interface optimization of “AI Jiusi 2.0”, and showcases the performance of the new version “AI Jiusi” in ancient Chinese language knowledge and language ability. **[Result/ Conclusion]** The newly upgraded “AI Jiusi 2.0” demonstrates significant advantages in the understanding of ancient Chinese texts and ancient Chinese knowledge question-answering, and has already acquired a certain ability to understand ancient Chinese characters (oracle bone script, bronze script) images, thereby making a due contribution to the advancement of ancient Chinese large language models.
Keywords: AI Jiusi2.0, Ancient Chinese, Large Language Model, Multimodal

1 引言

2024 年 10 月 28 日，习近平总书记在主持中共中央政治局第十七次集体学习时强调，要“探索文化和科技融合的有效机制，实现文化建设数字化赋能、信息化转型，把文化资源优势转化为文化发展优势”。仅一周后，习近平总书记来到湖北考察时再次指出，系统推进历史文化遗产保护传承和活化利用，让中华文明瑰宝永续留存、泽惠后人，激励人们不断增强民族自豪感和自信心。古代汉语作为中华传统文化的核心载体，承载了几千年来中华民族的智慧结晶与文化基因，不仅是沟通古今的桥梁，更是铸牢文化根基、赓续中华文脉的关键所在。然而，古代汉语的独特性和复杂性，使其在古籍保护利用和文化传承中面临诸多挑战。

随着生成式人工智能（AIGC）时代的到来，各类大模型如雨后春笋般涌现，并由最初仅能处理单一文本模态的大语言模型（Large Language Model，LLM），逐步升级为能够同时处理文本、图像、语音和视频等多模态数据的多模态大语言模型（Multimodal Large Language Model，MLLM），在全球范围内掀起了 AI 领域的“诸侯争霸”（如下图 1[1]）。在通用领域，以 GPT-4V[2]、Gemini-1.5[3]、Qwen-VL[4]、VisualGLM[5]等为代表的国内外多模态大语言模型，在文本或图像的理解与生成、跨模态检索、多模态问答等领域展现了强大的能力，也为古代汉语的教学、研究和应用带来了新的机遇。但因古汉语领域专业数据和知识规模的限制，这些模型在处理古汉语信息处理下游任务、回答古汉语领域相关专业知识问题时，表现却不尽人意。

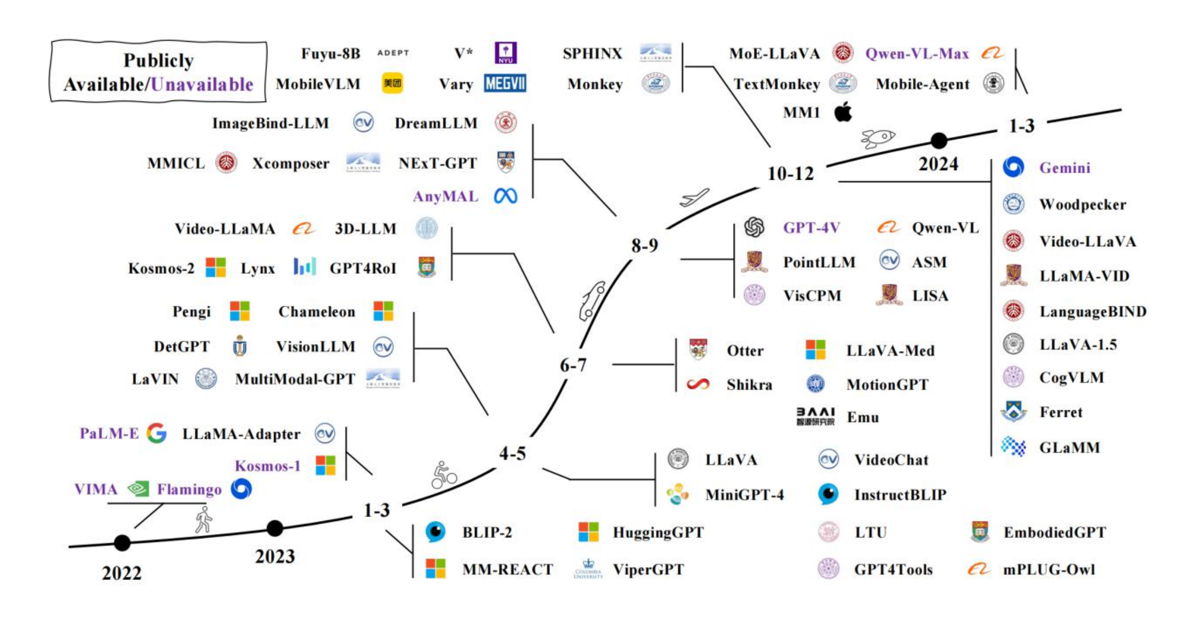


图 1：代表性多模态大语言模型发展时间线（Yin S, Fu C, Zhao S, et al,2023.）

为解决这一问题，国内多所高校几乎同时展开了面向古代汉语专业领域的大语言模型研发，并取得了阶段性成果。北京师范大学于 2023 年 11 月启动古汉语文本理解大语言模型“AI 太炎 1.0”的内测，并于 2024 年 8 月发布“AI 太炎 2.0”公众版[6]。南京农业大学于 2023 年 12 月发布古籍大语言模型“AI 荀子 1.0”，并于 2024 年 5 月推出“AI 荀子 2.0”[7]。这些模型目前主要聚焦于提升句读标点、文白翻译、实体识别等古汉语信息处理任务的性能，且以单一文本模态信息处理为主，在大语言模型的知识理解与问答交互能力，以及多模态信息处理方面，尚待进一步深入探究。

基于此，我们于 2024 年元旦开启国内首个兼具古汉语知识问答和理解应用能力的古汉语认知大语言模型“AI 九思 1.0”的内测，并由此开启了多模态古汉语大语言模型的研发探索之路[8]。经过近一年的努力，研发团队现全新推出既掌握古汉语专业知识，又兼备古汉语应用能力，并支持多模态数据处理的古汉语多模态大语言模型——“AI 九思 2.0”，为多模态古汉语大语言模型的发展抛砖引玉。

2 AI 九思 2.0 模型简介

“AI 九思 2.0”是由华中科技大学人文学院汉籍数字化实验室、铸牢中华民族共同体意识研究基地刘根辉教授团队在“AI 九思 1.0”的基础上，全面升级、全新研发的多模态古代汉语大语言模型。该模型在古汉语的文本理解能力、专业知识问答等方面取得了显著突破，同时新增古文字图片理解的多模态数据处理能力，进一步拓展了其应用广度和深度。

“AI 九思 2.0”不仅能够更加高效准确地完成智能句读标点、词法分析、文白翻译、实体识别、通假识别及典故解析等古汉语文本理解任务，同时更加全面而深入地掌握了文字学、音韵学、训诂学、方言学、目录学、版本学、校勘学等古汉语多领域的专业知识，还能够实现对甲骨文、金文图片的智能识别与形、音、义释解。其功能构成如图 2 所示。



图 2：“AI 九思 2.0”功能构成

2.1 “AI 九思 2.0”数据集构建

高质量数据集的构建与优化是提升“AI 九思 2.0”性能的基础。如图 3 所示，为确保模型能够更加精准把握古汉语的特殊性和复杂性，我们通过多源采集、系统整理、严格标注、多轮评估的四级流程，建立了系统化的数据集构建体系。

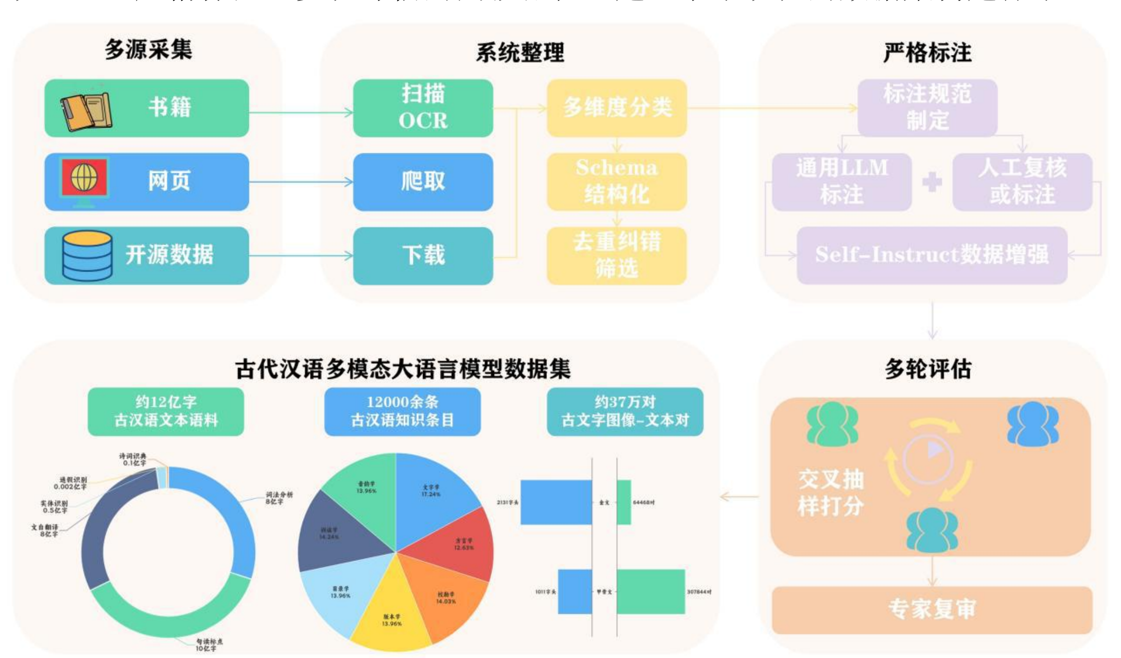


图 3：“AI 九思 2.0”数据集构建流程

（1）多源采集

在原始数据收集环节，我们以权威版本文献为核心，数字资源为补充，开源数据为辅助，构建了一套多源互补的数据采集体系。具体而言，以古汉语领域 150 余种经典权威纸质文献为核心，通过扫描、OCR 识别、人工精校等，获取了文献精校电子版本。这些文献涵盖权威出版机构出版的古籍译注本、古汉语经典教材、专业参考书，以及古文字研究领域的经典论著、工具书等。同时，我们广泛收集古汉语领域公开网站上的相关数据，并合理利用现有开源数据，进一步拓展数据集的广度与深度。这种多源采集体系，全方位保障了数据的权威性、系统性、多样性与全面性，为模型训练输送了丰富的原始数据。

（2）系统整理

在完成数据收集后，我们采用自动化与人工相结合的方式对原始数据进行了系统化整理，为预训练阶段的无监督数据准备及后续指令微调数据集的构建奠定了基础。首先，根据文本理解、知识问答、图像理解等不同任务类型，为原始数据定义了相应的元数据，并依此对数据进行多维度分类。以文本理解数据为例，我们定义了数据来源、作者、时代、部类、繁简、句读情况、数据用途（如句读、翻译、实体识别等）等元数据，并通过结构化 Schema 存储这些元数据，实现了数据的标准化和结构化。最后，利用自然语言处理技术对数据进行了去重和纠错处理，确保了数据的纯净性和数据结构的一致性。

（3）严格标注

在数据整理的基础上，我们组建了一支由中文专业特别是古汉语研究领域的本硕博学生为核心的专业标注团队，进行指令微调数据集的标注工作。标注团队依据我们制定的《古代汉语大语言模型数据集构建标注规范 V2.0》，采用自动化标注与人工标注相结合的方式开展工作：对于句读标点、文白翻译等任务，数据来源于纸质文献，进行人工复查、入库；对于标注尚未一致的数据，采用 Few-Shot Prompting[9]（少量样本提示）方式，调用三种不同通用大模型 API 进行粗标注，再对标注一致的数据进行人工复查、入库；对于甲骨文、金文图像理解数据及知识问答数据，则完全由人工标注后再复查、入库。

同时，为了提升数据集规模和泛化性，我们还将初步标注完的高质量数据作为种子指令数据，基于 Self-Instruct 技术（自我指令生成技术）[10]，利用通用大语言模型自动生成多样且意图类似的 Prompt 指令，并基于此引导模型解析古汉语知识问答数据集中已有的古汉语知识内容，从而自动生成与这些知识内容逻辑相关的新问答对。

（4）多轮评估

标注完成后，我们进行了多轮评估，以确保数据集的高质量与可靠性。评估过程采用标注团队交叉抽样打分与专家复审相结合的方式。首先，标注团队相互交叉，从已标注的数据中随机抽取样本，依据上述标注规范中的数据质量要求，从准确性、一致性、完整性等多个维度对样本进行逐项评分；之后，由项目专家对抽样数据进行复审，重点评估标注数据的学术准确性，进一步修正了数据集中可能存在的偏差或错误，确保了数据集的学术严谨性。

基于上述操作流程，我们成功构建了一个覆盖全面、质量可靠、标注规范的古汉语多模态大语言模型数据集。目前，数据集已包含超过 12 亿字的古汉语精校文本、1.2 万余条古汉语知识条目及 37 万余对古文字“图片—释文”对，为“AI 九思 2.0”的训练提供了坚实的数据支撑。



图 4：“AI 九思 2.0”数据集构成

如图 4 所示，我们对古汉语原始文本数据进行系统整理，经过去重、纠错、筛选、分类、结构化等环节，最终获得超 12 亿字规模的海量语料。这些语料在

继续预训练阶段全部投入使用，在有监督微调阶段则因任务适配性改写，语料会有不同程度的损耗。

2.2 “AI 九思 2.0”算力升级

算力的提升是推动“AI 九思 2.0”实现高效训练和性能突破的核心要素。为了应对多模态大语言模型训练中的巨大计算需求，并进一步缩短模型训练周期，我们依托华中科技大学的高性能计算公共服务平台，对计算资源进行了全面升级。该平台是由华中科技大学统一规划建设的校级公共计算服务平台，它以超级计算、人工智能计算、大数据处理等软硬件为基础，面向全校各单位提供大规模科学计算和数据处理服务。目前，平台拥有 369 个计算节点，近 2 万个 CPU 核心，104 块 GPU 卡，可用存储空间超过 4PB，理论峰值算力达到 2.5Pflops 以上。

这些先进的计算资源不仅提供了强大的并行计算能力，还大幅增加了内存容量和存储空间，确保能够高效处理复杂任务。这使得“AI 九思 2.0”的训练效率得到了显著提升，同时也为模型的多模态融合训练提供了可能。通过高效完成对海量古汉语文本及图像数据的学习和理解，我们能够快速迭代模型版本，不断优化模型性能。

2.3 “AI 九思 2.0”模型训练

在数据与算力的双重加持下，“AI 九思 2.0”的模型训练进入了全新的智能进阶阶段。我们以阿里云通义千问 Qwen2.5-7B[11]和 Qwen2-VL-7B[12]为基座模型，采用了两阶段训练策略，结合多模态训练和 RAG（Retrieval-Augmented Generation，检索增强生成）技术[13]，确保模型在古汉语领域的深度理解与多模态数据处理能力上实现新的突破。

（1）两阶段训练：从基础到精专，促进模型能力再进阶

“AI 九思 2.0”先后进行了继续预训练（Continued Pretraining）和有监督微调（Supervised Fine-tuning）两阶段的训练，结合 DeepSpeed ZeRO 2 技术[14]，优化了显存占用，提高了资源利用率。

在继续预训练阶段，“AI 九思 2.0”依托我们构建的高质量古汉语数据集，进行了大规模的无监督学习。通过对超过 12 亿字的古汉语精校文本进行深度学习，模型逐步掌握了复杂多样的古汉语文字的构形特征、文本的语法结构和语义特征以及知识内涵，大大增强了其在处理古汉语多模态数据时的泛化能力。

在有监督微调阶段，我们设计了一系列针对古代汉语文本理解、知识问答、文字图像问答的下游任务，通过引入自行构建的高质量古汉语指令微调数据集，进一步对模型进行了有针对性的微调。微调过程中，我们采用了 LoRA（Low-Rank Adaptation，低秩自适应）技术[15]，通过在模型中引入低秩矩阵来优化参数更新，不仅减少了所需的训练时间和计算成本，还有效防止了过拟合，使得模型在目标任务上的表现更加优异。

（2）多模态训练：从文本到图像，实现跨模态深度理解

“AI 九思 2.0”的一大突破在于其多模态数据处理能力。我们特别设计了多模态训练流程，将古文字图片（如甲骨文、金文）和对应释文文本进行联合对齐训练，使模型能够理解两者之间的关联性。例如，模型可以根据甲骨文字图片识别出其对应的现代汉字，并进一步解析文字的形、音、义。

（3）RAG 增强：检索增强生成，缓解模型幻觉更精准

为了进一步提升“AI 九思 2.0”的性能，我们还引入了 RAG（Retrieval-Augmented Generation，检索增强生成）技术，通过结合检索与生成两种能力，使模型能够在回答问题时，首先从我们构建的古汉语知识库中检索相关信息，之后

结合检索结果与自身的语言生成能力，生成精准且符合上下文的答案。这种检索增强的生成方式，不仅提高了输出答案的准确性，还可确保模型能够回答更加复杂和专业的古汉语问题，同时有效缓解模型在生成过程中可能出现的“幻觉”问题。其检索增强生成流程如图 5 所示。

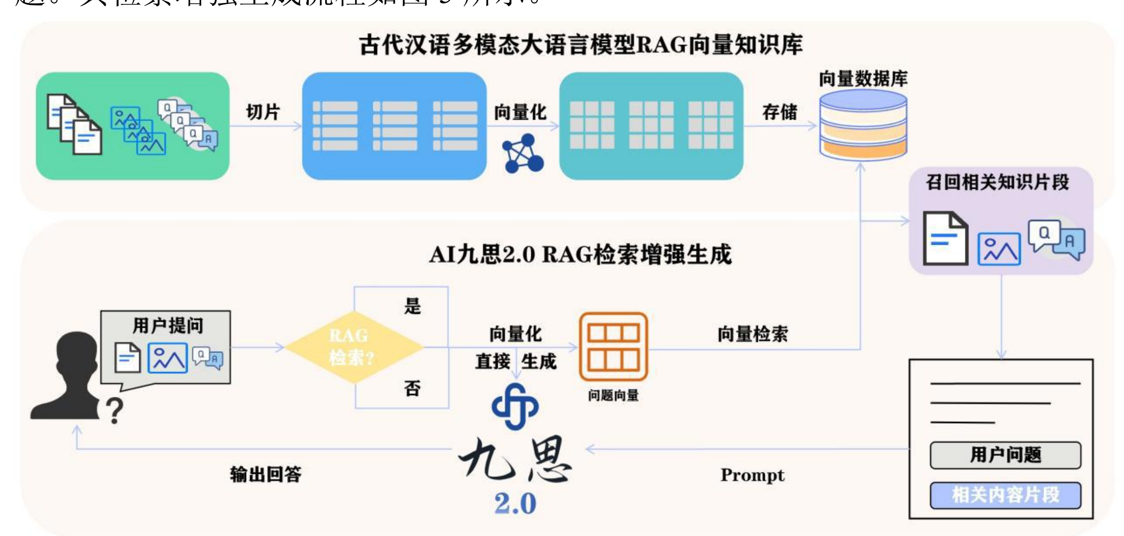


图 5：“AI 九思 2.0”RAG 检索增强生成流程

2.4 “AI 九思 2.0”界面优化

为了让用户能够更加便捷、高效地使用“AI 九思 2.0”，我们基于 THChat UI[16]进行二次适配开发，在界面设计和用户交互上进行了全面优化，力求为用户提供更加美观、友好的使用体验。

（1）界面设计更美观

“AI 九思 2.0”采用了全新的界面设计，整体风格简洁大方，功能布局清晰明了。主页设计遵循现代简约风格，每个功能模块都配有直观的文字说明，用户可以轻松了解和使用所需功能。同时，用户可以根据个人喜好，在如云白、水墨黑和古典雅多种主题之间自由切换，满足不同用户的视觉偏好。无论是现代简约风格，还是古典雅致的界面，都能为用户带来舒适的使用体验。

（2）用户交互更友好

“AI 九思 2.0”在交互设计上进一步简化了操作流程，提供了灵活的会话管理功能。用户可以轻松创建、删除会话，随时切换不同的对话场景，确保交互体验更加灵活自如。同时，系统支持多轮对话，用户可以与“AI 九思 2.0”进行连续的对话交流，系统能够根据上下文提供更加精准的回答，极大提升了交互的流畅性和自然度。

此外，“AI 九思 2.0”还支持多模态交互，用户不仅可以通过普通文本与系统对话，还可以直接上传古文字（如甲骨文、金文）的图片，系统可自动识别并释解图像中的内容。同时，系统全面支持简体中文和繁体中文的输入与输出，用户可以根据自己的习惯选择使用繁简字体，系统能够自动识别并处理，确保不同用户群体的使用需求都能够得到满足。

3 “AI 九思 2.0”新功能展示

“AI 九思 2.0”在全面提升 1.0 版本句读标点、词法分析、实体识别、通假字识别、文白翻译等文本理解与文字、音韵、训诂、目录、版本、校勘等知识问答各项功能的基础上，新增甲金文字识解、诗词典故识解、方言学知识问答三大

功能。

3.1 甲金文字识解

多模态图文并解，可智能识别甲骨文、金文单字图片，并基于古文字学家研究成果，给用户提供尽可能准确的形、音、义释解，助力古文字研究与学习。

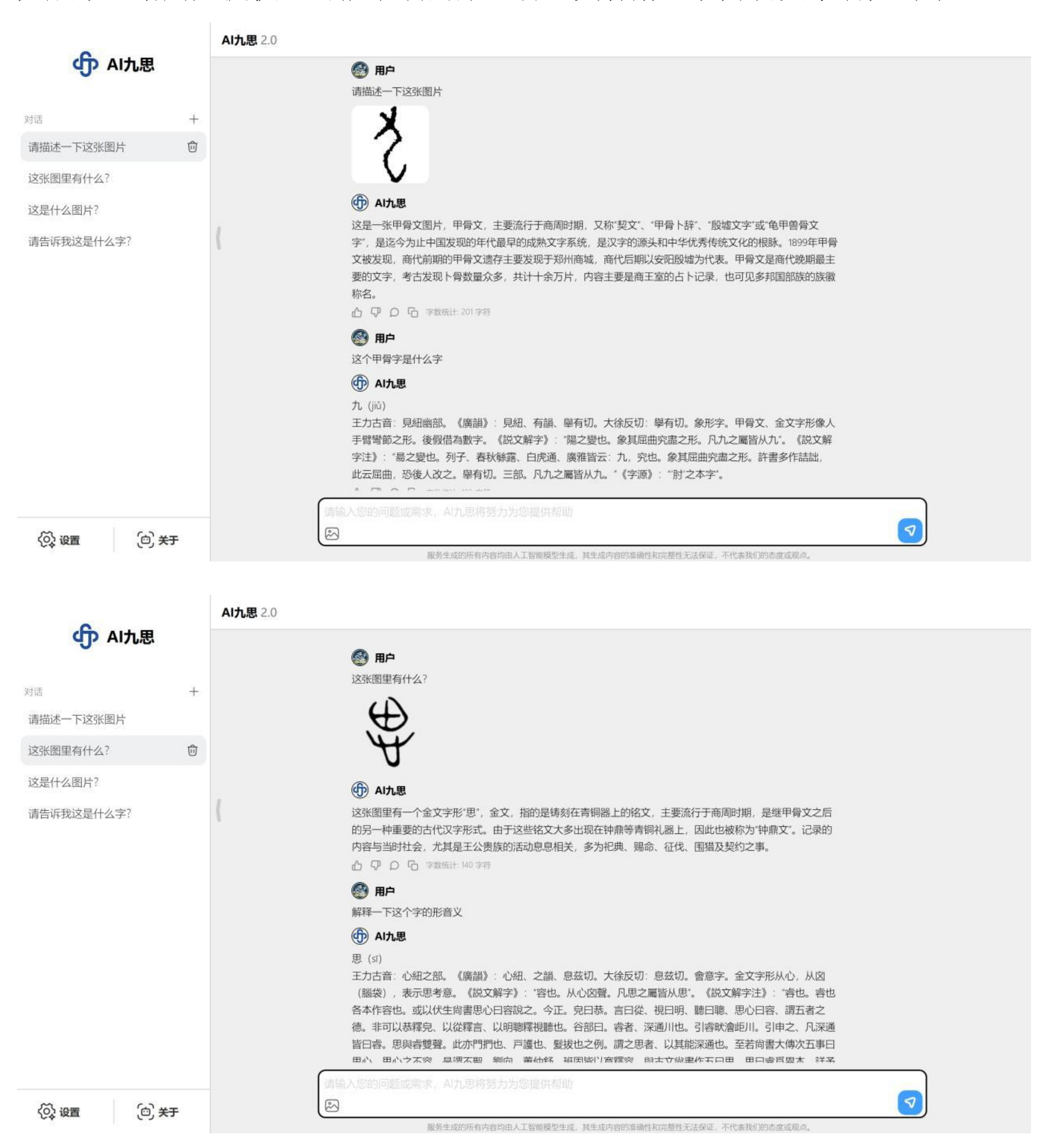


图 6、图 7：“AI 九思 2.0”甲金文字识解：图中分别展示了“AI 九思 2.0”识别并释解甲骨文字形“九”、金文字形“思”的图片

3.2 诗词典故识解

深度挖掘唐诗、宋词等古典文学作品中的典故，实现自动识别与解析，帮助用户更深入地体味古典诗词之美。

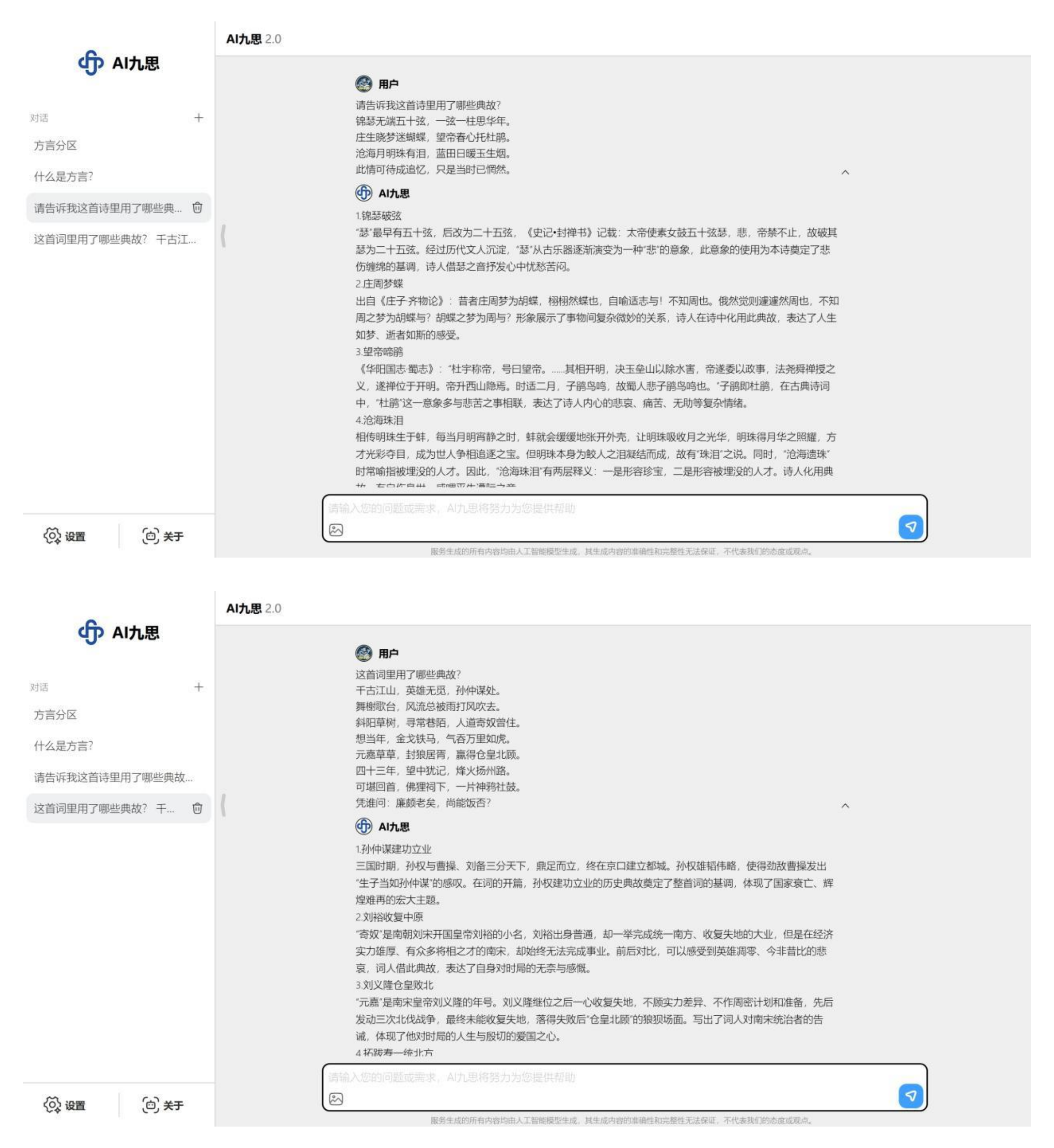


图 8、图 9：“AI 九思 2.0”诗词典故识解：图中分别展示了“AI 九思 2.0”识别并释解唐代诗人李商隐的《锦瑟》、宋代词人辛弃疾的《永遇乐·京口北固亭怀古》所用典故

3.3 方言学知识问答

构建涵盖古代汉语方言学知识的专业知识库，能够回答用户关于方言演变、方言分区、方言特点等方面的专业问题，为用户提供便捷的方言学知识获取渠道。

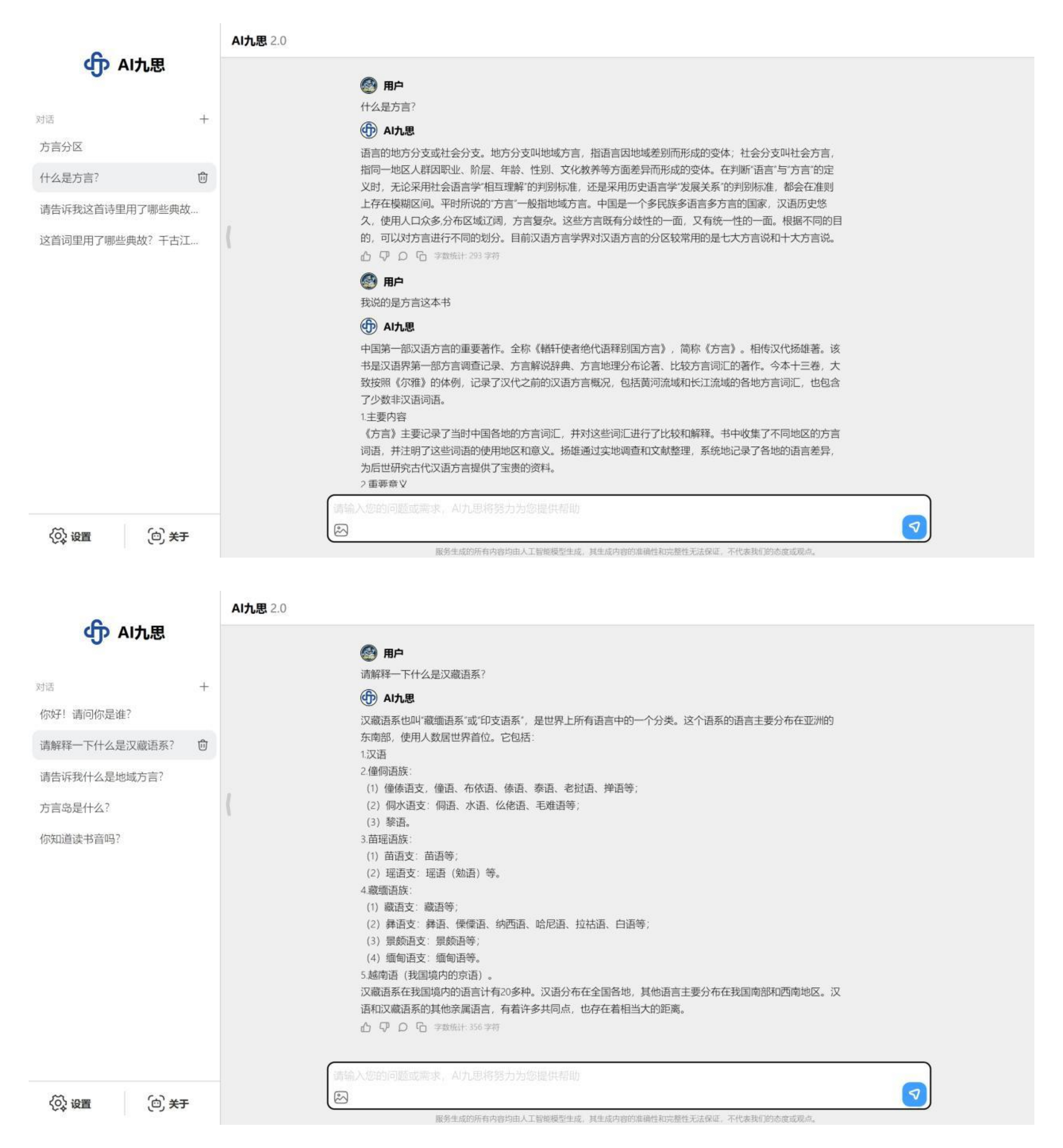


图 10、图 11：“AI 九思 2.0”方言学知识问答：图中分别展示了“AI 九思 2.0”回答“什么是方言”、“介绍扬雄《方言》”、“什么是汉藏语系”等问题

4 “AI 九思”未来展望

在各界的支持和鼓励下，“AI 九思 1.0”在古代汉语大语言模型的研发方面实现了“从 0 到 1”的突破。“AI 九思 2.0”则是研发团队进一步探索和推动古代汉语大语言模型向多模态方向发展的全新力作。“AI 九思”将秉承“君子九思”的精神，不忘初心，踔厉奋发，勇毅前行，踵事增华。

我们期待，“AI 九思 2.0”的推出能够推动古代汉语大语言模型向更高层次演进，逐步实现从“大语言模型”到“大模型”乃至“智能体”的跨越式发展。未来，“AI 九思”将继续为“探求古汉语之菁华、赓续古汉语之绝学、传承古汉语之美”贡献力量，助力中华优秀传统文化的继承与弘扬！

让我们共同见证古代汉语数智化信息处理的未来，携手开启古汉语研究的新篇章！

参考文献：

[1] Yin S, Fu C, Zhao S, et al. A survey on multimodal large language models[J]. arXiv preprint arXiv: 2306.13549,2023.

[2] Yang Z, Li L, Lin K, et al. The Dawn of LMMs: Preliminary Explorations with GPT-4V(ision) [J]. 2023.

[3] Team G, Georgiev P, Lei V I, et al. Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context[J]. arXiv preprint arXiv:2403.05530,2024.

[4] Bai J, Bai S, Yang S, et al. Qwen-vl: A versatile vision-language model for understanding, localization, text reading, and beyond[J]. arXiv preprint arXiv:2308.12966,2023,1(2):3.

[5] 智谱 AI. VisualGLM-6B[EB/OL]. https://github.com/THUDM/VisualGLM-6B?tab=readme-ov-file, 2024-08-23.

[6] 李绅,胡韧奋,王立军.古汉语大语言模型的构建及应用研究[J].语言战略研究,2024,9(05):22-33.

[7] 王东波课题组. 荀子古籍大模型 [EB/OL].https://github.com/Xunzi-LLM-of-Chinese-classics/XunziALLM, 2024-11-30.

[8] 刘金柱,王锦绣,罗捷春,李志芳,袁方,刘艺溶,刘根辉等. AI 九思：用大语言模型焕新古汉语之美[DB/OL]. (2025-01-23)[2025-01-25].https://chinaxiv.org/abs/202501.00212v1.

[9] Wei J, Tay Y, Bommasani R, et al. Emergent abilities of large language models[J]. arXiv preprint arXiv:2206.07682,2022.

[10] Wang Y, Kordi Y, Mishra S, et al. Self-instruct: Aligning language models with self-generated instructions[J]. arXiv preprint arXiv:2212.10560,2022.

[11] Team Q. Qwen2.5: A party of foundation models[J]. Qwen (Sept.2024).url:https://qwenlm.github.io/blog/qwen2,2024,5.

[12] Wang P, Bai S, Tan S, et al. Qwen2-vl: Enhancing vision-language model's perception of the world at any resolution[J]. arXiv preprint arXiv:2409.12191,2024.

[13] Lewis P, Perez E, Piktus A, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks [J]. Advances in Neural Information Processing Systems, 2020,33:9459-9474.

[14] Rajbhandari S, Rasley J, Ruwase O, et al. Zero: Memory optimizations toward training trillion parameter models[C]//SC20: International Conference for High Performance Computing, Networking, Storage and Analysis. IEEE, 2020:1-16.

[15] Hu E J, Shen Y, Wallis P, et al. Lora: Low-rank adaptation of large language models[J]. arXiv preprint arXiv:2106.09685, 2021.

[16] Unagi-cq. LM 的 Web 会话管理方案 THChatUIv2[EB/OL].https://github.com/Unagi-cq/THChatUI,2025-01-13.

（通信作者：刘根辉　E-mail：chnlab@mail.hust.edu.cn）

作者贡献声明：

大语言模型“AI 九思 2.0”是在项目负责人刘根辉教授带领下的研发团队所有成员集体智慧的结晶。

刘金柱：完成全部实验流程，论文初稿撰写；

王锦绣：负责数据集构建，论文修改；

罗捷春，李志芳，袁方，余静静，龚丹，谢雨霏，罗婉滢，郑苏楠，陈旷心，贺心雨，张润哲，夏婉婷，谢佳延，吕佳源，吕萍，余乐妍，郑诗铭，王金柳，刘艺溶，杨纯，张曼丽，吴翊嘉，余锁湘，汪靓：数据集构建；

刘根辉：论文整体框架设计，论文定稿，研究思路制定。